Effects of Sidetone Amplification on Vocal Function During Telecommunication*

*,†,*Nicole E. Tomassi, ‡M. Eugenia Castro, ‡Lauren Timmons Sund, †Manuel E. Díaz-Cádiz, †,§Daniel P. Buckley†, and *,†,§,||Cara E. Stepp, *†§||*Boston, Massachusetts, and* ‡*Los Angeles, California*

Abstract: Purpose. Society has become increasingly dependent on telecommunication, which has been shown to negatively impact vocal function. This study explores the use of sidetone regulation during audio-visual communication as one potential technique to alleviate the effects of telecommunication on the voice.

Method. The speech acoustics of 18 participants with typical voices were measured during conversational tasks during three conditions of sidetone amplification: baseline (no sidetone amplification), low sidetone amplification, and high sidetone amplification. Vocal intensity, vocal quality (estimated using acoustic measures of the low-high ratio and the smoothed cepstral peak prominence), and self-perceived vocal effort were used to measure the impacts of sidetone amplification on vocal function.

Results. Compared to baseline, there were statistically significant decreases in vocal intensity and increases in low-high ratio in the high level of sidetone amplification condition. Changes in these measures were not significantly correlated. When asked to rank conditions based on their perceived vocal effort, participants most often ranked the high level of sidetone amplification as least effortful; however, the visual-analog ratings of vocal effort were not significantly different between conditions. The smoothed cepstral peak prominence did not change with varying levels of sidetone amplification.

Conclusions. Vocal intensity decreased with high levels of sidetone amplification. High levels of sidetone amplification also resulted in increases in the low-high ratio, which were shown to be more than just a byproduct of decreased vocal intensity. The impact of sidetone amplification on vocal effort was less clear, but results suggested that participants generally decreased their vocal effort with increased levels of sidetone amplification. This was a preliminary study and future work is warranted in a population of participants with voice complaints and in a more noisy, realistic environments.

Key words: Telecommunication−Speech production−Vocal function−Live-mic monitoring−Auditory feedback.

INTRODUCTION

The use of technology to communicate remotely has become a new standard in various sectors of society, ranging from work and education to health and leisure. While there are benefits of convenience for this mode of communication, research has shown that frequent use of telecommunication impacts vocal function. Typical speakers report vocal symptoms such as increased vocal effort with increasing use of both audio and audiovisual telecommunication platforms for their work e.g., [1,2]. Often, these speakers are required to spend long hours telecommunicating with few breaks and poor vocal hygiene. [3,4] Studies have observed short-term changes in voice production during the use of telecommunication platforms, finding increases in both vocal intensity and speaker-reported vocal effort when compared to in-person communication. [1,5] The results are consistent with retrospective evidence suggesting that individuals who occupationally use telecommunication and speech recognition software may be at increased risk for symptoms of muscle tension dysphonia. [6-8] Whether short-term or chronic, voice changes during telecommunication are likely due, in part, to disruptions in the sensorimotor monitoring mechanisms that speakers use to maintain intelligibility during communication, as evidenced by phenomena such as the Lombard effect and sidetone regulation.

The Lombard effect describes the tendency of speakers to make involuntary adjustments of vocal intensity when background noise is increased.[9] Increases in background noise are also accompanied by rises in fundamental frequency, flattening of the voice spectral slope, and lengthening of signal duration, especially for vowels. [10] These vocal changes are most prominent in conversational tasks, relative to simple reading or other non-communicative phonation e.g., [11,12], indicating that the speaker's goal is intelligible communication. Such involuntary adjustments are critical during telecommunication, as compared to in-person communication. Remote interactions pose added challenges due to the lack of beneficial verbal and non-verbal cues that foster intelligibility when in-person.

Accepted for publication March 23, 2021.

*Address correspondence and reprint requests to Nicole E. Tomassi, Graduate Program for Neuroscience, Boston University, Boston, MA

Disclosure: Cara Stepp has received consulting fees from Altec, Inc./Delsys, Inc., companies focused on developing and commercializing technologies related to human movement. Stepp's interests were reviewed and are managed by Boston University in accordance with their conflict of interest policies. All other authors have declared that no competing interests existed at the time of publication.

From the *Graduate Program for Neuroscience, Boston University, Boston, Massachusetts; †Department of Speech, Language, and Hearing Sciences, Boston University, Boston, Massachusetts; ‡USC Voice Center, Caruso Department of Otolaryngology − Head and Neck Surgery, University of Southern California, Los Angeles, California; §Department of Otolaryngology − Head and Neck Surgery, Boston University School of Medicine, Boston, Massachusetts; and the ||Department of Biomedical Engineering, Boston University, Boston, Massachusetts.

E-mail: ntomassi@bu.edu

Journal of Voice, Vol. ■■, No. ■■, pp. ■■−■■
0892-1997
Published by Elsevier Inc. on behalf of The Voice Foundation.
https://doi.org/10.1016/j.jvoice.2021.03.027

Sidetone regulation, is a related phenomenon in which speakers use their own vocal feedback to regulate their vocal amplitude. Much like the Lombard effect, sidetone regulation is prominent during conversational tasks. [11-14] Sidetone regulation and the Lombard effect have been hypothesized to result from the same underlying processes that aim to maintain intelligibility during conversation. [9,12] That is, during conversation the speaker must compensate for any changes in the signal-to-noise ratio in order to remain intelligible. It is often the case during telecommunication that a high background noise or reduced sidetone warrants the speaker to perceive their intelligibility as low, causing them to make changes in their voice production in attempts to increase the signal-to-noise ratio of conversation. Thus, increasing sidetone amplification may have the potential to offset the negative vocal effects of telecommunication. For instance, when participants are provided feedback of their own voice through headphones, their vocal intensity decreases. [12,13,15-19] These studies have demonstrated this effect in an experimental environment, using repetitive speech tasks in ideal acoustic settings; however, this effect has also been observed in more ecologically valid environments. For example, a 2003 study showed similar findings, observing how teachers change their voice production when given self-feedback during a day of teaching. [20] Additionally, an earlier study added sidetone amplification into live business calls to understand sidetone regulation in a setting that represented a real-life environment. [14] However, it is not known whether these changes in vocal intensity transfer to changes in voice quality and vocal effort. Control in the level of sidetone amplification can be achieved with low-latency live-mic monitoring technology, which provides real-time playback of one's voice through a headset. Use of such technology could potentially mitigate the vocal risks associated with use of telecommunication platforms

Current research has primarily focused on the negative effects and probable risks due to the increasing use of telecommunication; however, far fewer studies have attempted to determine techniques to mitigate such risks. Thus, this study aims to determine the potential therapeutic effects of sidetone amplification on vocal function during audio-visual telecommunication. Two levels of sidetone amplification were compared to a baseline condition with no sidetone amplification to assess the effects on vocal intensity, vocal quality (estimated using acoustic measures of the low-high ratio and the smoothed cepstral peak prominence), and self-perceived vocal effort. We hypothesized that sidetone amplification during telecommunication would have an effect on voice production of speakers with typical voices such that, when compared to the baseline condition, higher levels of sidetone amplification would result in lower vocal intensity, higher low-high ratios, lower smoothed cepstral peak prominence values, and lower self-perceived vocal effort.

METHODS

Participants

Eighteen cisgender, non-smoking, native speakers of English (eight males and ten females; M_{age} = 21.0 years SD_{age} = 2.53 years) participated in the study. The participants self-reported no history of speech, language, hearing, or neurological disorders and provided written consent as per the Boston University Institutional Review Board. All participants passed a hearing screening at levels of 25 dB HL at the frequencies of 125, 250, 500, 1000, 2000, 4000, and 8000 Hz. Participants reported their daily average time spent on telecommunication platforms. Eight participants reported daily averages between 2-4 hours and ten participants reported daily averages of more than 5 hours. All participants completed the Voice-Related Quality of Life (V-RQOL) questionnaire with raw scores ranging from 10 to 26 (M_{VRQoL} = 13.33, SD_{VRQoL} = 4.73), indicating that participants did not perceive their voice as having a significant impact on their day-to-day activities.

Procedure

Per COVID-19 protocol, all participants wore masks upon arrival and experimenters were in full personalized protective equipment for the duration of the study. Once participants were seated alone in a sound-attenuated booth, they were then instructed to remove their mask. The experimenter remained outside the booth for the duration of testing. All instructions were communicated to the participant by the experimenter through an Amazon Echo Show (Amazon.com, Inc, Seattle, WA). Participants wore a SHURE MX153 omnidirectional microphone (SHURE, Niles, Illinois) angled 45 degrees lateral from the midline and positioned 7 cm away from the corner of the mouth. Sennheiser HD 280 Pro over-ear headphones (Sennheiser Electronic GmbH & Co, Wennebostel, Germany) were used for playback capabilities. A Knowles BU 21771 accelerometer (Knowles Acoustics, Itasca, Illinois) was placed at the notch of the neck, secured with adhesive tape, and used to record neck surface vibrations. To calibrate speech recordings an electrolarynx positioned 7 cm from the participant's microphone, in line with the mouth, was used to generate output. The sound pressure level in dB SPL was measured using a Galaxy CM-150 sound pressure level meter (Galaxy Audio, Wichita, Kansas), which was held in close proximity to the microphone. SONAR Artist acoustic software was used to collect acoustic recordings on a desktop computer at a sampling rate of 44,100 Hz.

A single experimenter (the first author) virtually interacted with all participants and wore a SHURE dynamic headset microphone (SHURE, Niles, Illinois). To capture voice-related vibrations, a K&K Hot Spot transducer (K&K Sound, Coos Bay, Oregon) was fixed on the surface of the experimenter's neck. Both the microphone and transducer signals were recorded with a LS-10 Linear PCM digital handheld audio recorder (Olympus Corporation of the

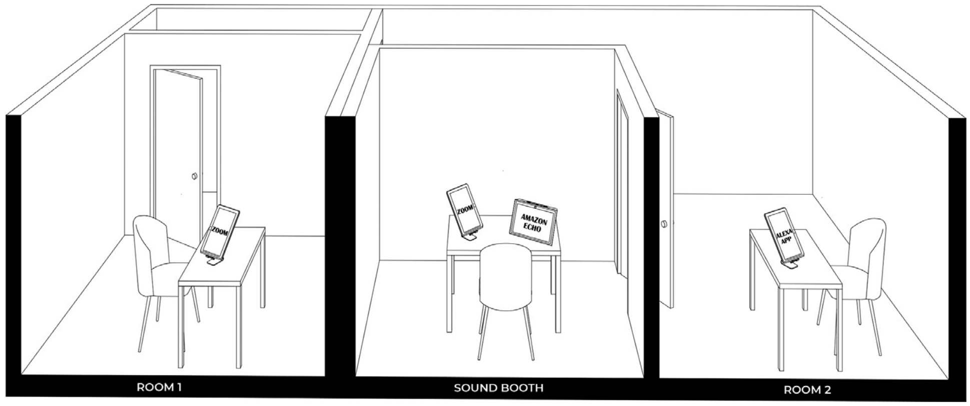

FIGURE 1. Schematic representation of the experimental setup. Room 2 contained the sound booth and experimental equipment; this was where the participant was located for the duration of the experiment. The participant sat in front of two devices set up for audio-visual communication. The Amazon Echo device was set up for instructions throughout the experiment while the device using Zoom was set up for the specific experimental conditions. Likewise, the device located on the desk outside of the sound booth was set up with the Amazon Alexa application for the experimenter to give the participant instructions. Room 1 was where the experimenter was located during each condition. Here, the experimenter sat in front of the device using Zoom to communicate with the participant during experimental conditions.

Americas, Center Valley, Pennsylvania) at a sampling rate of 44,100 Hz.

Three portable Samsung Galaxy Tab E tablets (Samsung Electronics, Seoul, South Korea) with volumes set to 80% were used in this study. Two were used for telecommunication, one by the experimenter and one by the participant, during an audio-visual conference call using Zoom Video Communications. The third was used for the experimenter to connect to the Amazon echo located inside the sound booth to provide set-up instructions to the participant. Figure 1 illustrates a schematic representation of the experimental setup. Prior to the start of the experiment, the tablets used for the Zoom conferences were stationed on tablet stands with the screens positioned 45 degrees upwards and were not moved for the duration of the experiment. One was placed within the sound booth, 32 inches away from the participant, and the other was placed in a small conference room nearby. This tablet was also stationed approximately 32 inches away from the experimenter.

In order to understand the effects of sidetone amplification on voice production, the participants completed three conversational tasks with varying levels of sidetone amplification. During baseline, only the experimenter's voice was emitted through the headphones. The participants could hear themselves via the bone-conducted and attenuated (32 dB) air-conducted natural feedback. The other two tasks provided feedback through the headphones that varied in terms of the gain applied to the sidetone feedback. The participants' voices were digitized and transmitted through their headphones with a delay of 11ms, which is below the 16-26 ms threshold that has been demonstrated to be perceived by a listener as an echo. [21] Thus, this allowed the participants to hear both themselves and the experimenter as the same time. A Brüel & Kjær sound level meter, microphone preamplifier, and headphone coupler (Brüel & Kjær, Nærum, Denmark) were used for calibration, such that participants heard their voice amplified to -3 dB and -9 dB relative to their microphone signal in these two tasks.

The experimenter and participant completed three versions of a collaborative communicative task identical to the ones performed in Tracy, et al. [1] and adapted from the scenarios used by Seita, et al. [22] The study consisted of tasks approximately 10 minutes in duration with differing scenarios and level of sidetone amplification. For each task, designed to simulate an extemporaneous conversation, the participant and experimenter collaborated on determining 10 items of significance if stranded in a specific environment and then ranked the items from most to least important. The specific environments ranged between lost at sea, stuck on the moon, and stranded in a desert. These environments corresponded to participant sidetone amplification levels of baseline, low (-9 dB), and high (-3 dB) respectively. The environmental scenarios with the matched sidetone amplification level were considered conditions; conditions were counterbalanced across participants. The participants were given a clipboard, paper, and pen to write notes during the conversations.

Participants were asked to take a five-minute voice break between conditions. During this time, the participants rated their self-perceived vocal effort on a 100-mm visual analog scale [VAS]. [23] Vocal effort was defined for the participants as "how easy or difficult it is to talk in terms of how much effort, strain, discomfort, and/or fatigue you perceive when using your voice, independent of your volume." The scale was labeled "No Effort" at 0 mm and "Max Effort" at 100 mm. Once all three conditions and their associated VAS ratings were completed, participants were instructed to rank-order the conditions in terms of vocal effort level from most (1) to least (3) vocal effort exerted.

Data Analysis

Initial signal processing was implemented using a custom MATLAB script (MATLAB 2018, Mathworks). Acoustic signals of the participants were calibrated to the electrolarynx tone. Times were extracted in which only the participant was speaking in order to obtain their isolated speech acoustics. The microphone and accelerometer signals of both the participant and experimenter were time-aligned. The microphone signal was used to analyze the properties of the voice whereas the accelerometer signals were used solely to identify when the participant and experimenter were speaking. A threshold was applied to the envelope of each participant's accelerometer signal to remove pauses. Similarly, a threshold was applied to the envelope of the experimenter's accelerometer signals to remove times during which the experimenter was speaking. In order to ensure appropriate pause and voice segmentation, threshold values were manually extracted via visual and auditory inspection by the first author. The outcome of this processing was a continuous signal of speech solely from the participant with total signal lengths averaging 248 s (range: 85—567 s, SD: 110 s).

For each condition, each participant's mean vocal intensity in dB SPL was calculated in MATLAB software (MATLAB 2018, Mathworks). Likewise, each participant's low-high (LH) ratio and smoothed cepstral peak prominence (CPPS) were computed for each condition using PRAAT acoustic software.[24]

Statistical Analysis

To determine the effect of the sidetone amplification level (baseline, low, high) on the four measured outcomes (vocal intensity, LH ratio, CPPS, vocal effort), four repeated measures analyses of variance (ANOVAs) were used with condition as a factor. An alpha value of 0.05 or less was used for significance testing. For outcome measures for which the condition was a significant contributor, effect sizes were calculated using a squared partial curvilinear correlation (η_p^2) designated as small (~0 .01), medium (~ 0.09), or large (> 0.25), as described in Witte and Witte. [25] Post-hoc analyses were performed on any outcome measures for which the condition was a significant contributor using Tukey pairwise comparisons. To determine if there was an association between condition and the vocal effort rank-orders assigned by participants at the end of the experiment, a Chi square test was performed using the likelihood ratio test statistic. The likelihood ratio chi square test statistic is based on the ratio of observed to expected frequencies of occurrence. For this study, we expect the likelihood ratio to provide insight on how likely it is that a participant would rank their vocal effort in a certain order based on the condition. We hypothesized that, with increasing levels of side tone amplification, the participants would rank order their vocal effort in the opposite direction, i.e. higher level of side tone amplification resulting in a lower rank of vocal effort An effect size was calculated using Cramer's V with four degrees of freedom and designated as small (~0.05), medium (~0.15), or large (>0.25), as described in Cohen. [26]

Based on the primary study results, additional explanatory analyses were performed to determine the relationship between changes in voice quality (estimated by LH ratio) and vocal intensity. Per participant, differences in LH ratio and vocal intensity were compared between 1) low and baseline and 2) high and baseline using Pearson's correlation analysis.

RESULTS

Average values normalized per participant to the baseline condition for the four measured outcomes are shown in Figure 2. The results of the ANOVAs and associated effect sizes are listed in Table 1.

The results from the ANOVA indicated that condition was a statistically significant factor on vocal intensity with a large effect size (Table 1). The post-hoc analysis revealed a statistically significant decrease in vocal intensity between the high and low sidetone amplification (p_{adj} = 0.025) and between the high sidetone amplification level and baseline (P_{adj} = 0.019), but no difference between the low sidetone amplification level and baseline. The mean vocal intensity was 82.4 dB SPL (SD = 3.4 dB) during baseline, 82.4 dB SPL (SD = 3.1 dB) during the low sidetone amplification level, and 81.4 dB SPL (SD = 3.1 dB) during the high sidetone amplification level.

Condition was also found to be a statistically significant factor on LH ratio with a large effect size (Table 1). Post-hoc analysis described a statistically significant difference between the low and high sidetone amplification levels (P_{adj} = 0.001) and the high sidetone amplification level and baseline (P_{adj} = 0.003), but no difference between the low sidetone amplification level and baseline. The mean LH ratio was 25.1 dB (SD = 4.0 dB) during baseline, 24.8 dB (SD = 4.7 dB) during the low sidetone amplification level, and 27.1 dB (SD = 2.9 dB) during the high sidetone amplification level.

Condition was not found to be a statistically significant factor for CPPS or self-perceived vocal effort. The mean CPPS was 11.2 dB (SD = 1.2 dB) during baseline, 11.1 dB (SD = 1.2 dB) during the low sidetone amplification level, and 11.2 dB (SD = 1.1 dB) during the high sidetone amplification level.

Although not statistically significant, participants tended to decrease their vocal effort with increasing sidetone amplification levels, with an average vocal effort of 20.8 (SD = 16.0) during baseline, 18.0 (SD = 14.7) during the low sidetone amplification level, and 16.0 (SD = 15.2) during the high sidetone amplification level.

The Chi square analysis revealed a statistically significant association (P = 0.049) with a large effect size (V = 0.42) between condition and how participants rank-ordered the trials based on vocal effort. Of the trials with the high sidetone amplification level, 50% were ranked as least effortful (3); of the baseline trials, 50% were ranked most difficult

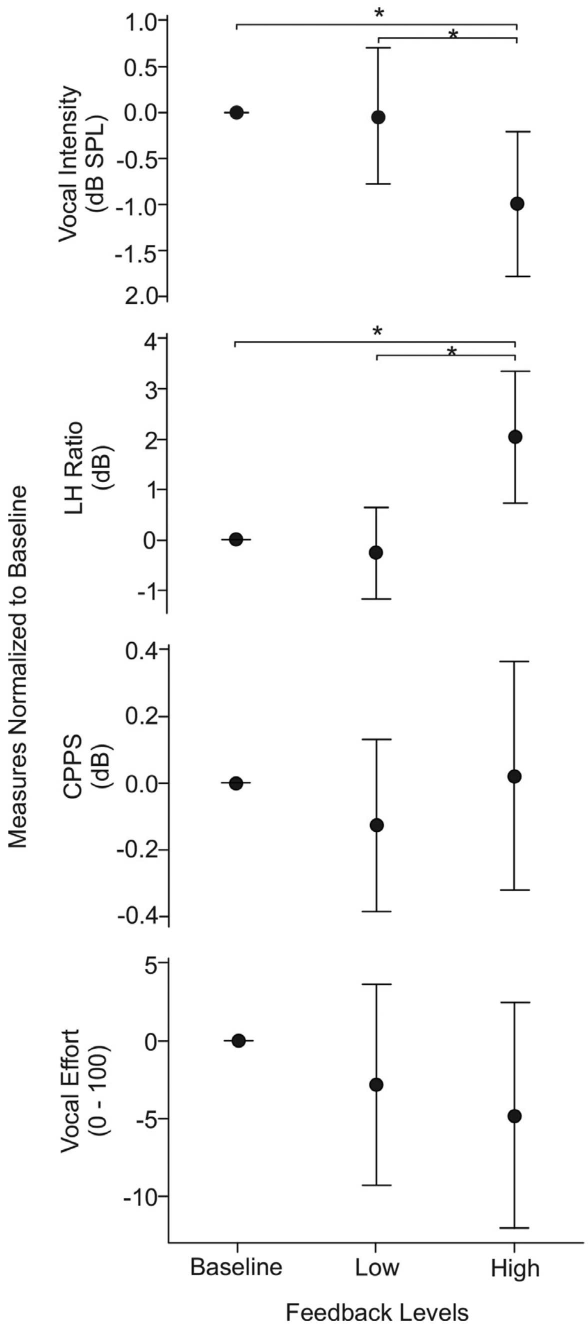

FIGURE 2. Mean and 95% confidence intervals relative to the baseline condition for vocal intensity, low-high (LH) ratio, smoothed cepstral peak prominence (CPPS), and vocal effort for the three sidetone amplification conditions—baseline, low (-9 dB), and high (-3dB). Asterisks denote statistical significance at an alpha level of 0.05 ($P_{adj} < 0.05$) between conditions.

(1); and of the trials with the low sidetone amplification level, 50% were ranked in the middle, as shown in Figure 3. This result suggests that participants were more likely to provide lower vocal effort rankings for conditions with higher levels of sidetone amplification.

The explanatory analyses on the relationship between LH ratio and vocal intensity did not result in statistically significant correlations. Specifically, the correlation analysis of the differences in LH ratio and vocal intensity between the low amplification level and baseline indicated a p = 0.636 and r = -0.12. The relationship between LH ratio and vocal intensity changes remained weak when examining the differences between the high amplification level and baseline, with a P = 0.512 and r = -0.16.

DISCUSSION

The purpose of this study was to determine the effects of sidetone amplification on vocal function during telecommunication. We hypothesized that participants would decrease their vocal intensity with increasing levels of sidetone amplification. In terms of acoustic estimates of vocal quality, we hypothesized that participants' LH ratios would increase and that their CPPS values would decrease with increased levels of sidetone amplification. Additionally, we hypothesized that participants would report decreased levels of vocal effort with increasing levels of sidetone amplification. Our hypotheses were supported by changes in the measures of vocal intensity and LH ratio, and by the ranked-order of self-perceived vocal effort.

As hypothesized, participants decreased their vocal intensity with increasing levels of sidetone amplification. This finding is consistent with the literature regarding sidetone regulation. [12,13,15-19] As sidetone amplification is increased, vocal intensity is often decreased. Our findings expand the work of McKown and Emling [14], showing that this effect generalizes from in-person communication, to telephone communication and here to audio-visual telecommunication. These adjustments in vocal intensity may be attempts from the speaker to remain intelligible during conversation across altered acoustic environments.

Based on our acoustic outcomes, vocal quality changed as a function of sidetone amplification. Specifically, participants phonated with a higher LH ratio as the level of sidetone amplification increased. Thus, there was more low-frequency or less high-frequency energy in the vocal spectrum with higher sidetone amplification. It is possible that the increases in LH ratio seen with sidetone amplification could be due to changes in high frequency noise. However, increased high-frequency energy would be aharmonic, which would have resulted in simultaneous decreases in CPPS, which was not seen in our study. Therefore, the more plausible physiological explanation for the increased LH ratio as the level of sidetone amplification increased, is that less medial compression of the vocal folds was used to produce vocalization (consistent with a less "pressed" voice). Reduced medial compression allowed for a relative increase in airflow, which results in a vibratory cycle with a shorter closed phase. This phonation pattern is often associated with a more prominent first harmonic in the low-frequency bands of the voice source spectrum relative to the high frequencies and may be captured by the LH ratio. [27] These changes could occur as a result of the noted changes in vocal intensity. However, a comparison of changes in LH ratio and vocal intensity indicated weak relationships. These

TABLE 1.
Results of analysis of variance models

Factor	Df	F-Value	*P*-value	Effect size	Qualitative Effect Size
		Vocal Intensity			
Sidetone Amplification Condition	2	5.25	0.010	0.24	Large
		LH Ratio			
Sidetone Amplification Condition	2	9.79	<0.001	0.37	Large
		CPPS			
Sidetone Amplification Condition	2	0.69	0.509	-	-
		Vocal Effort			
Sidetone Amplification Condition	2	1.22	0.307	-	-

Effect sizes and interpretations are only provided for significant effects ($P < 0.05$). Dashes indicate nonsignificant findings. *df* = degrees of freedom; CPPS = smoothed cepstral peak prominence; LH ratio = low-high ratio.

findings suggest that changes in vocal quality, measured by LH ratio, are not solely a byproduct of changes in vocal intensity.

Regardless of the physiological basis, LH ratio was significantly higher when speakers were exposed to higher sidetone amplification, yet CPPS was unchanged. Prior research has found that CPPS is a promising measure of dysphonia, correlating strongly with perceptions of severity, roughness, hoarseness, and strain. [28-30] Although spectral tilt measures have been shown to be correlated to breathiness in dysphonic speakers, the perceptual significance of measures of spectral tilt in typical voices is unclear. [31,32] Thus, from a clinical perspective, a higher LH ratio, as found in this study under higher sidetone amplification, may indicate improved vocal quality, but more research must be performed to support this claim since CPPS was not impacted by condition.

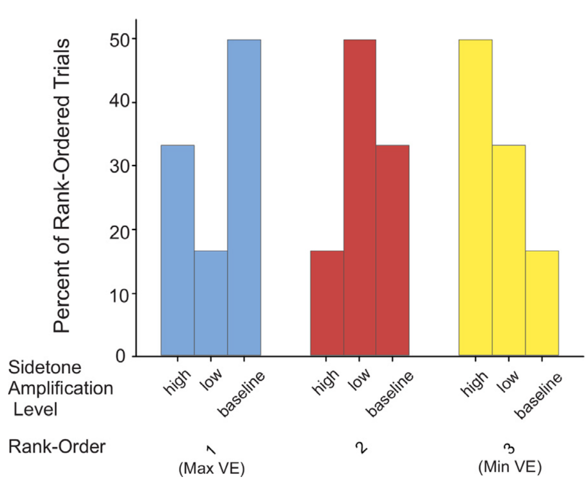

FIGURE 3. The percentage of vocal effort (VE) rank-orders assigned by participants at the end of the experiment for each condition. A statistically significant association ($P = 0.049$) was found between condition and vocal effort rank-orders, with a large effect size (V = 0.42). Thus, participants were more likely to provide a lower ranking with increased levels of sidetone amplification.

The impact of sidetone amplification on speakers' perceptions of vocal effort was less clear. When asked to rate vocal effort on a VAS, participants reported decreased levels of vocal effort with increased levels of sidetone amplification on average; yet this relationship was not supported statistically. However, when asked to rank-order the vocal effort required for each condition at the end of the experiment, participant's responses were statistically significantly different across the conditions. One reason for this disparity could be the large degree of variability across participants in reporting vocal effort. This variability was most prominent in the baseline and high sidetone amplification level conditions. Because vocal effort is a perceptual measurement, it is not uncommon for responses to be variable. For example, Tracy, et al. [1]studied the impact of communication modality on voice production and found that vocal effort among all communication modalities had a high degree of variability. Another possible cause of the weak effect of sidetone amplification on vocal effort could be the sample composition. Participants were young adults with typical voices and perceived vocal effort changes may be more useful when studying individuals with voice complaints or voice disorders. An additional cause of the weak effect of sidetone amplification on vocal effort could be the short task duration. The length of each condition (10 minutes) may not have been enough time to cause measurable changes in vocal effort. However, the cumulative effect of the use of increased vocal intensity over the course of a day of telecommunication is likely to lead to increased vocal effort, as seen in vocal loading tasks.[33]

This is a preliminary study of sidetone amplification as an intervention during telecommunication and thus has several limitations. One possible limitation is that the levels of sidetone amplification studied may not have captured a large enough range to elicit more robust differences between conditions. The levels were chosen empirically based on initial pilot testing in which amplification levels were altered based on the comfort and preferences of laboratory staff. In the future, a larger range of amplification levels could be used to determine if these affect the measured outcomes. In addition, this study was performed in ideal acoustic conditions;

it is possible that in a more realistic environment with background noise, the measured outcomes may have been altered. Another limitation of this study was that it was conducted 10 months into a global pandemic, requiring experiments to be as remote as possible. Thus, this study used only acoustic measures in order to be non-invasive, socially distanced, and safe. Future studies employing aerodynamic measures could provide more insight about the potential physiological bases of vocal change during sidetone amplification. Finally, only a select few acoustic measures were incorporated – vocal intensity, LH ratio, and CPPS. These measures were chosen because they are appropriate for running speech, which allowed us to implement a study design that was relatively ecologically valid. Despite these limitations, the use of sidetone amplification during telecommunication seems to be a promising way to improve some aspects of vocal function; instructions for clinicians and readers on how to implement this design for work and/or personal use can be found in the supplemental section.

CONCLUSIONS

This study explored the vocal effects of using sidetone amplification during audio-visual telecommunication. Sidetone amplification resulted in statistically significant decreases in vocal intensity and increases in LH ratio, an acoustic correlate of vocal quality. The impact of sidetone amplification on speakers' perceptions of vocal effort was less clear, but overall suggested that participants experienced less vocal effort with increased levels of amplification. These findings are tempered by the preliminary nature of the study. Future work in individuals with voice complaints in more realistic, noisy environments is warranted.

ACKNOWLEDGEMENTS

This work was supported by the National Institutes of Health grants R01DC015570 (CES), P50-DC015446 (REH), and T32DC013017 (CAM). The authors thank Hasini Weerathunge for aiding in calibration efforts and providing experimental training.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.jvoice.2021.03.027.

REFERENCES

[1]. Tracy L F, Segina R K, Cadiz MD, et al. The impact of communication modality on voice production. *J Speech Lang Hear Res*. 2020;63(9):2913–2920.
[2]. Jones K, Sigmon J, Hock L, et al. Prevalence and risk factors for voice problems among telemarketers. *Arch Otolaryngol Head Neck Surg.*. 2002;128(5):571–577.
[3]. Fuentes-López E, Fuente A, Contreras KV. Inadequate vocal hygiene habits associated with the presence of self-reported voice symptoms in telemarketers. *Logoped Phoniatr Vocol*. 2019;44(3):105–114.
[4]. Johns-Fiedler H, van Mersbergen M. The prevalence of voice disorders in 911 emergency telecommunicators. *J Voice*. 2015;29(3):389.e1–389.e10.
[5]. Kisenwether JS, Anson D. Cell Yell!: health risks in Telehealth. *Perspectives of the ASHA Special Interest Groups*. 2019;4(3):538–541.
[6]. Olson DE, Cruz R M, Izdebski K, et al. Muscle tension dysphonia in patients who use computerized speech recognition systems. *Ear Nose Throat J*. 2004;83(3):195–198.
[7]. Lehto L, Laaksonen L, Vilkman E, et al. Occupational voice complaints and objective acoustic measurements—do they correlate? *Logoped Phoniatr Vocol*. 2006;31(4):147–152.
[8]. Haxer M J, Guinn LW, Hogikyan ND. Use of speech recognition software: a vocal endurance test for the new millennium? *J Voice*. 2001;15(2):231–236.
[9]. Lombard E. Le signe de l'elevation de la voix. *Ann Mal de L'Oreille et du Larynx*. 1911:101–119.
[10]. Zollinger S A, Brumm H. The Lombard effect. *Curr Biol*. Aug 23 2011;21(16):R614–R615.
[11]. Abeyesekera A, Adams SG. Effects of altered intensity feedback on speech in healthy speakers. *Can Acoust*. 2020;48(3).
[12]. Lane H, Tranel B. The Lombard sign and the role of hearing in speech. *J Speech Lang Hear Res*. 1971;14(4):677–709.
[13]. Lane H, Catania A C, Stevens SS. Voice level: Autophonic scale, perceived loudness, and effects of sidetone. *J Acoust Soc Am*. 1961;33(2):160–167.
[14]. McKown F, Emling J. A system of effective transmission data for rating telephone circuits. *Bell Syst Tech*. 1933;12(3):331–346.
[15]. Siegel GM, Pick Jr HL. Auditory feedback in the regulation of voice. *J Acoust Soc Am*. 1974;56(5):1618–1624.
[16]. Garnier M, Henrich N, Dubois D. Influence of sound immersion and communicative interaction on the Lombard effect. *JSLHR*. 2010.
[17]. Laukkanen A M, Mickelson NP, Laitala M, et al. Effects of HearFones on speaking and singing voice quality. *J Voice*. Dec 2004;18(4):475–487.
[18]. Siegel GM, Pick HL, Olsen MG, et al. Auditory feedback on the regulation of vocal intensity of preschool children. *Dev Psychol*. 1976;12(3):255.
[19]. Chang-Yit R, Pick Jr HL, Siegel GM. Reliability of sidetone amplification effect in vocal intensity. *J Commun Disord*. Dec 1975;8(4):317–324.
[20]. Jónsdóttir V, Laukkanen AM, Siikki I. Changes in teachers' voice quality during a working day with and without electric sound amplification. *Folia Phoniatr Logop*. 2003;55(5):267–280.
[21]. Lezzoum N, Gagnon G, Voix J. Echo threshold between passive and electro-acoustic transmission paths in digital hearing protection devices. *Int J Ind Ergonom*. 2016;53:372–379.
[22]. Seita M, Albusays K, Kafle S, et al. Behavioral changes in speakers who are automatically captioned in meetings with deaf or hard-of-hearing peers. In: *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. 201868–80.
[23]. McKenna V S, Stepp C E. The relationship between acoustical and perceptual measures of vocal effort. *J Acoust Soc Am*. 2018;144(3):1643–1658.
[24]. Boersma P. Praat, a system for doing phonetics by computer. *Glot Int*. 2001;5(9):341–345.
[25]. Witte RS, Witte JS, Wiley. *Statistics*. Ed.. 9 ed. 2010.
[26]. Cohen J. *Statistical power analysis for the behavioral sciences*. Academic press; 2013.
[27]. Zhang Z. Mechanics of human voice production and control. *J Acoust Soc Am*. 2016;140(4):2614–2635.
28. Heman-Ackah YD, Sataloff RT, Laureyns G, et al. Quantifying the cepstral peak prominence, a measure of dysphonia. *J Voice*. 2014;28(6):783–788.
[29]. Maryn Y, Corthals P, Van Cauwenberge P, et al. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice*. 2010;24(5):540–555.

[30]. Heman-Ackah YD, Michael DD, Goding Jr GS. The relationship between cepstral peak prominence and selected parameters of dysphonia. *J Voice*. 2002;16(1):20–27.

[31]. Hillenbrand J, Cleveland RA, Erickson RL. Acoustic correlates of breathy vocal quality. *J Speech Lang Hear R*. 1994;37(4): 769–778.

[32]. Hillenbrand, J, Houde RA. Acoustic correlates of breathy vocal quality: dysphonic voices and continuous speech. *J Speech Lang Hear*. 1996;39(2):311–321.

[33]. Chang A, Karnell MP. Perceived phonatory effort and phonation threshold pressure across a prolonged voice loading task: a study of vocal fatigue. *Journal of Voice*. 2004;18(4):454–466.